Causal Model of a Health Services System

by James G. Anderson

Path analysis is used to construct a causal model of the health services system serving the state of New Mexico. The model includes a network specifying the causal relationships among a set of social, demographic, and economic variables hypothesized to be related to the health status of the population; a set of mathematical equations that permit prediction of the effects of changes in the values of any one variable on all other variables in the model; and estimates of path coefficients based on U.S. Census data and vital statistics. The model is used to predict both direct and indirect effects on health status of changes in population structure resulting from natural causes or from the intervention of health programs.

Effective program planning and evaluation necessitates more than the accumulation of time series data concerning the health of the population and the functioning of the health services delivery system. In order to evaluate the effectiveness of any program, it is first necessary to be able to establish and demonstrate causal relationships between program inputs and the consequences of that program for the target population.

To date most attempts to evaluate the effectiveness of health services delivery systems have gone no further than elementary attempts to conceptualize indicators of state variables which at best lend themselves to summary and description but allow for no further inferences. Few of these efforts have attempted to develop models that reliably depict the interrelations among variables characterizing the target population, operating characteristics of the delivery system, and the effects on the population of the care delivery [1].

The nature of health services makes the development of such models especially difficult. Adequate operational definitions of effectiveness are elusive and appropriate data are often unavailable [2–5]. Moreover, as Packer [6] has pointed out, elements of the health services system such as doctors, hospitals, and neighborhood health clinics typically operate in an independent fashion, with many aspects of their operation fixed for long periods of time. At the same time they are subject to changing patterns of demand and utilization as a result of advances in medical science, changing patterns of medical insurance coverage, and population shifts, among other factors.

Supported in part by NCHSRD contract HSM-110-69-243 and in part by a Faculty Research Grant from the Purdue Research Foundation.

Address communications and requests for reprints to James G. Anderson, Ph.D., Dept. of Sociology, Purdue University, Lafayette, In. 47907.

To further complicate matters, health service systems are stochastic [7] and frequently behave in ways that are opposite to expectations [8]. Forrester [9] has remarked that the short-term response of a complex system to policy changes is apt to be in the opposite direction from its long-term response, which appears to be largely the result of feedback. In health service delivery systems, not only will the social, demographic, and economic characteristics of the target population affect demand and utilization, but the delivery system, if effective, can also be expected to affect the social, demographic, and economic characteristics of the population. In fact, the expectation of such feedback effects provided the rationale for Section 202 of the Appalachian Regional Development Act, which authorizes HEW grants for multicounty demonstration health projects. The reasoning used to carry the bill through Congress emphasized the importance of health resources in attracting industry to the region [10].

Feedback effects are also involved in decisions regarding the location of health facilities. Such decisions are normally based on assessments of current patterns of utilization, but the location of the new facility in turn will affect the pattern of utilization. Studies by Turner [11], Cardwell et al. [12], Shannon et al. [13], Weiss and Greenlick [14], and Weiss et al. [15] have demonstrated that location affects the aggregate demand for health services as well as the method by which individuals enter the health system, with the effect of distance being largely mediated by social and economic characteristics of the population such as race, education, and income.

Migration, too, may have a profound effect on patterns of demand and utilization of services, an effect resulting from changes in the social, economic, and health characteristics of the population served. It has been repeatedly demonstrated that changes in welfare programs providing more attractive benefits in one area than another can result, in a relatively short period of time, in significant changes in the composition of the population served, with attendant increased demands on the delivery system. Any program aimed at removing barriers to the utilization of health services by low-income minority groups must consider the complex repercussions of such a program on social and demographic processes such as these in order to make decisions regarding location of planned facilities, future staffing of facilities, and the like.

Almost all the attempts to date to develop models of health service systems or subsystems have involved microlevel models that trace an individual's movement through the health system. A number of publications have reviewed the early attempts to apply operations research models to health systems [16–20]. From studies involving the storage and distribution of physical resources such as nursing unit supplies [21], food [22, 23], linens [24], and pharmacy items [25, 26], researchers turned their attention to the utilization of facilities. One approach was to consider expected demand for hospitalization as a function of economic and demographic factors [27–29].

Other researchers utilized the Poisson process to describe utilization of health services [30] and developed queueing models to describe the admission process [31, 32], the flow of patients through outpatient clinics [33, 34], and the operation of a physician's office appointment system [35, 36]. Confronted with the restrictions imposed by classical queueing theory, Bailey [37], with Welch [38], was among the first to use Monte Carlo techniques to model hospital outpatient departments. Subsequent models have been developed that take advantage of special simulation languages and modern computing equipment [39–41].

Most of the more recent attempts to model health service systems, while more comprehensive in scope, still involve microlevel models that trace an individual's movement through some health subsystem. Thus Rosenhead [42] simulated the drug prescription process of a physician by randomly generating patients and hypothetically assigning them to a physician. Colley et al. [43] developed simulation models for nursing homes that generate estimates of resource levels and operating costs for different levels of simulated patient care. Fetter and Thompson [39] simulated, among other subsystems, the flow of patients through a maternity suite, an outpatient clinic, and a surgical pavilion in a hospital, using the computer simulation language, SIMSCRIPT.

Comprehensive attempts to simulate a community health service system are reported by Gentry et al. [44], by Kennedy [45], and by Kennedy and Woodside [46]. Milly and Pocinki [47] describe a model that will simulate the progress of patients through a delivery system involving facilities, personnel, and equipment grouped into service-treatment stations.

The major justification for this microlevel approach is that it leads to a health service systems model that accurately portrays the real-world health system in enough detail so that the system's performance and operating costs in response to changes in inputs, facilities, health manpower, and services can be estimated [39, 43, 44, 47]. But such measurements of health system performance and operating costs fail to take into account the complex interrelations between the health system and the social structure of the target population and thus permit consideration of only such things as operating costs, staffing levels, patient waiting time, and average service times by type of service when comparing alternative programs. The impact on the social, economic, and demographic composition of the population of demographic processes that affect the utilization of health services, such as migration and urbanization, are not included; nor is feedback such as the effect of changes in spatial distribution of facilities on patterns of utilization, or the effect of initial contact with and mode of entry into the health system on subsequent utilization of health services. Such effects need to be considered if the total effects of alternative programs are to be evaluated.

The model described in this article uses path analysis [48, 49] and causal inference [50, 51], based on ideas developed originally in biology [52–55] and economics [56, 57], to predict an index of health. These model building procedures begin with the specification of a causal structure involving a network of causal paths among variables.

This approach provides a means of constructing mathematical models that depict causal relationships among variables postulated to be related to the utilization and the effectiveness of a health service delivery system; they lead to the definition of a set of equations that correspond to the hypothesized causal processes occurring in the real world and the estimation of model parameters. From these equations it is possible to predict how a change in any one variable in the system affects the values of other variables in the system. The validity of the models can be evaluated from extant data and the models reformulated if necessary, or they can be used to predict the probable effect of changes in the population structure on health service requirements and to evaluate the effect of modifications in the organization and utilization of health service resources.

Path Analysis

Path analysis begins with the selection of a limited number of variables that are thought to characterize the social process under investigation. Social, economic, and demographic characteristics of the population have been shown to be important determinants of health status. Certain of these variables, such as the age composition of the population, affect health directly; others may have important indirect effects on health through their effects on variables more closely linked to health. Thus the educational level of the population has both direct effects on health and indirect effects through an income variable, which in turn mediates access to and utilization of health services, with consequences for health status.

The next step in the construction of a path model is the unambiguous causal ordering of variables on the basis of hypotheses concerning the direct and indirect effects of the variables on the health service delivery system and on the health status of the population. The 11 variables selected for this model are:

Exogenous (variation assumed caused by variables outside system)

X_1 Percent of labor force in agriculture 1968
X_2 Percent urban 1960
X_3 Percent Spanish-American or Mexican-American 1960
X_4 Percent nonwhite 1960

Endogenous

X_5 Net migration 1960–69
X_6 Median age 1960
X_7 Median education 1960
X_8 Percent unemployed 1968
X_9 Per capita income 1967

Health services index

X_{10} Hospital beds/population ratio 1968

Health status index

X_{11} Mortality from accidents, suicide, and cirrhosis of the liver 1968

A residual variable assumed to be uncorrelated with the other endogenous variables was also introduced, to account for the variation in each endogenous variable not accounted for by its relation with other variables in the system. The rationale for this selection of variables and the hypothesized interrelations are described below.

Health Status Index

The performance of any health service delivery system ultimately must be appraised in terms of its ability to improve the health of a target population. In the past, indexes based on mortality have been used as measures of levels of health [3]. More recently several authors have proposed measures based on both mortality and morbidity [2–5].

In order to create an index of health status for the purposes of this research, mortality data for New Mexico counties for the year 1968 were obtained from the New Mexico State Department of Health and Social Services. Causes of death were grouped into the following categories: heart disease, cancer, accidents, vascular lesions, influenza and pneumonia, diseases of early infancy, bronchopulmonic diseases, cirrhosis of the liver, diseases of the circulatory system, suicide, and all other causes. These causes of death were converted to death rates and a factor analysis was performed. Rates of mortality due to heart disease, cancer, vascular lesions, and bronchopulmonic diseases all loaded on one factor. A second factor appeared to contrast mortality due to diseases of early infancy with that from diseases of the circulatory system. A third factor reflected mortality due to influenza and pneumonia and to all other causes. Rates of mortality due to accidents, suicide, and cirrhosis of the liver all had high loadings on a fourth factor.

Since these last three causes of death have all been linked to sociocultural characteristics of the population, especially race and social class, they were combined into a single measure, which was used as an index of the health status of the population of each county in New Mexico. It is this measure that is used here to illustrate the model-building technique described, but a number of other mortality and morbidity measures have also been successfully used with the model, including the infant mortality rate; rates of mortality from influenza and pneumonia; heart disease, cancer, vascular lesions, and bronchopulmonic diseases; rates of morbidity from venereal diseases; and food- and water-borne diseases, bacterial and viral diseases, and acute respiratory diseases.

Health Services Variable

The variable most closely linked to health services in the model is the hospital beds/population ratio. Hospitals may be considered major providers of care in New Mexico, since the number of physicians per 100 000 population in the state is only one-half that of the U.S. population as a whole (1968 figures, 80 vs. 154). Such data as number of physician visits are not available for all New Mexico counties, so the beds/population ratio was selected as an index of health services available to the population.

Sociodemographic Variables

Both the facility measure and the health status index are next linked to several sociodemographic variables. The first of these measures, per capita income, is highly related to the health status of the population. Several publications from the National Center for Health Statistics adequately document the relationships among economic characteristics, health status, and health services utilization [58–61]. For example, the National Health Survey disclosed that 29 percent of the population with family incomes below $2000 suffer from chronic conditions that limit their usual activities, as against only 7.5 percent with family incomes above $7000. Moreover, the under-$2000 group on the average experienced more than twice the number of days of restricted activity per year as the latter group and had four and one-half times as many heart conditions, six times as much hypertension, arthritis, and rheumatism, more than eight times as many visual impairments, and six times as many mental and psychiatric disorders.

Despite these major differences in health, persons from low-income families underutilize health services. While 73 percent of those from families with incomes of $10 000 or more had visited a physician within the past year, only 64 percent from families below the poverty line ($3000 a year) had done so [62].

Income level of a community is also an important determinant of the health services available to the population. Consequently the hypothesized model also includes a link between per capita income and the hospital beds/population ratio. Income is itself the result of the overall state of the economy of an area, in particular the availability of job opportunities. Counties with significant amounts of unemployment or underemployment have lower per capita incomes. Moreover, health services are generally lacking in areas experiencing severe unemployment, while health problems are more prevalent among the members of such populations.

Education, too, is related to health both directly and indirectly, with direct effects of low educational level stemming from poor nutrition, sanitation, and personal hygiene and from lower frequency of preventive measures. Indirect effects result from the link between education and income, which in turn is related to the prevalence of chronic diseases, levels of disability, and health service utilization patterns.

The age composition of the population has important health implications. Data from the National Health Survey indicate that per capita physician visits are higher among persons under 5 years of age and those over 65 [63]. While acute illnesses occur most frequently in the under-5 group, chronic diseases account for many of the health problems of the over-65 group. This older group is, on the average, likely to be hospitalized for two weeks at a time, while the average length of stay in the hospital for persons under 45 is about half as long [64]. Moreover, causes of illness, mortality rates, and leading causes of death in a population are all highly related to its age structure.

The age structure also has important implications for the economic status of the population, with attendant indirect effects on health. For example, a high dependency ratio of young and old to those in the age group that makes up most of the labor force results in a major economic burden for the population. In addition, the age structure of the population affects levels of both education and income.

Population changes resulting from migration are also related to health status. Counties experiencing out-migration tend to have older populations with lower educational and income levels. Under these circumstances health services and the health of the population tend to deteriorate.

Finally, the model includes several exogenous variables. The percentage of the labor force engaged in agriculture is an important characteristic of the economic structure of a county. Agrarian-based counties on the whole are experiencing out-migration, which results in an older, poorly educated population. Per capita income in these areas is below that of more industrialized areas. At the same time the age structure of the population results in heavy demands on the limited health services available in the rural areas.

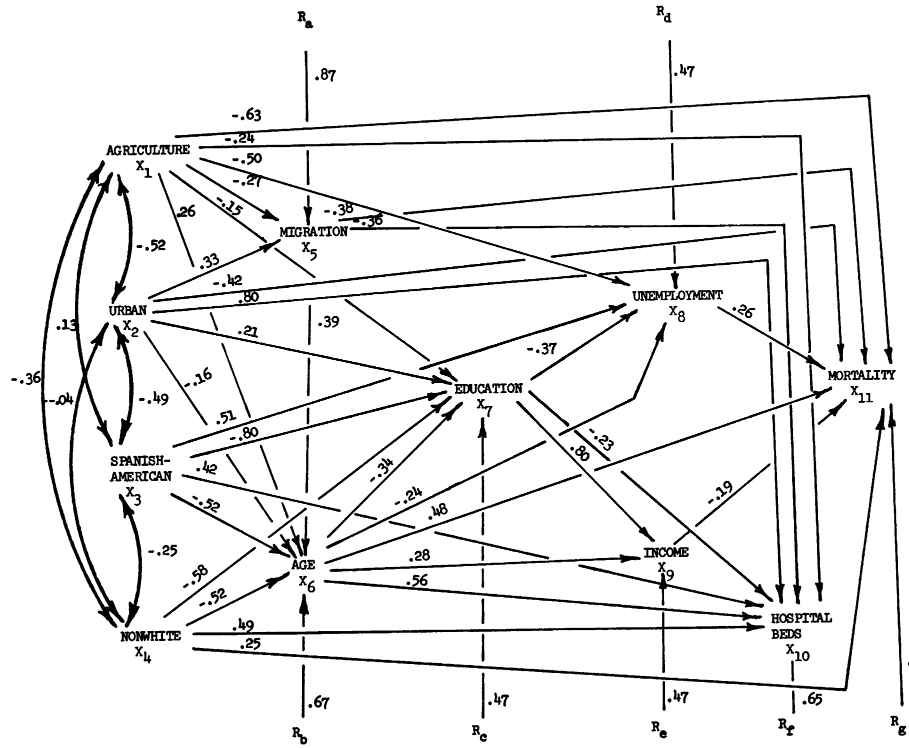
Urbanization is another significant factor affecting the level of health in a population. Increased urbanization of a population affects kinship relationships [65], the structure of the family [66], and the division of labor and level of technological development [67]. Health services are generally more available to urban populations. In addition to these indirect effects, there are significant rural-urban differences in morbidity and mortality [68].

The presence of large racial and minority ethnic groups in a county has profound effects on health as well as on the social and economic characteristics of the population. The model therefore includes two exogenous variables, percent Spanish-American (or Mexican-American) and percent nonwhite.

Richardson's studies [62] have shown that the nonwhite segment of the population underutilizes health services. For example, while white children under the age of 15 from families with incomes under $5000 on the average visit a physician 3.3 times per year, their nonwhite peers average only 1.9 visits. This differential obtains for every age and income category. The pattern of utilization differs also: while only 7.7 percent of visits by whites were to hospital clinics or emergency rooms, fully 26 percent of nonwhite visits occurred in this context.

On the whole, the nonwhite population is more urbanized than the white population in the United States. In 1960 over 70 percent of nonwhites lived in urban areas, as compared with only a little over 42 percent of whites. Educational levels, income levels, and a host of other variables connected with socioeconomic status also differentiate whites from nonwhites.

The Spanish-American population of the Southwest is similar in many respects. While this population was largely concentrated in urban areas by 1960, about one-half of those still residing in rural areas were engaged in agriculture. Median education levels for this group are below national levels. As a result a substantial number of Spanish-Americans had incomes of less than $3000 in 1959. Data from Colorado [69] indicate that Spanish-Americans more frequently die from accidents and from influenza and pneumonia, and less frequently from chronic conditions.



Path model relating social, demographic, and economic characteristics to health. Path coefficients and arrows for values less than .15 have been omitted to simplify the diagram.

Causal Model

The accompanying figure shows a model of the process by which the socioeconomic and demographic characteristics hypothetically affect the health status of the population. It depicts linear, additive, asymmetric relationships among the 11 variables selected to describe the structure and performance of a social system affected by a health and medical care delivery system. In the model, one-way arrows lead from each independent to each dependent variable in turn. As defined by Land [49, pp. 8–9], the path coefficients shown beside each of these arrows measure

> . . . the fraction of the standard deviation of the endogenous variable (with the appropriate sign) for which the designated variable is *directly* responsible in the sense of the fraction which would be found if this factor varies to the same extent as in the observed data while all other variables (including residual variables) are constant.

These path coefficients (standardized partial regression coefficients) are listed in full in Table 1 (overleaf); to simplify graphic presentation, only those paths with coefficients greater than .15 are shown in the figure. The structural equations from which the path coefficients are derived are given below. The curved lines with two-headed arrows connecting the exogenous variables in the figure represent unanalyzed correlation between these variables.

Data for each of the 32 counties in New Mexico were obtained from the U.S. Census, the New Mexico Bureau of Business Research, and the New Mexico State Department of Health and Social Services.

Structural Model

An important property of path diagrams is their isomorphism with the algebraic and statistical properties of the relationships postulated among the variables [48, 49]. The causal model depicted here can be written as a set of structural equations representing the hypothesized causal relationships that underlie the model, as follows:

$$
\begin{aligned}
X_5 &= p_{54}X_4 + p_{53}X_3 + p_{52}X_2 + p_{51}X_1 + p_{5a}R_a \\
X_6 &= p_{65}X_5 + p_{64}X_4 + p_{63}X_3 + p_{62}X_2 + p_{61}X_1 + p_{6b}R_b \\
X_7 &= p_{76}X_6 + p_{75}X_5 + p_{74}X_4 + p_{73}X_3 + p_{72}X_2 + p_{71}X_1 + p_{7c}R_c \\
X_8 &= p_{87}X_7 + p_{86}X_6 + p_{85}X_5 + p_{84}X_4 + p_{83}X_3 + p_{82}X_2 + p_{81}X_1 + p_{8d}R_d \\
X_9 &= p_{98}X_8 + p_{97}X_7 + p_{96}X_6 + p_{95}X_5 + p_{94}X_4 + p_{93}X_3 + p_{92}X_2 + p_{91}X_1 + p_{9e}R_e \\
X_{10} &= p_{10,9}X_9 + p_{10,8}X_8 + p_{10,7}X_7 + p_{10,6}X_6 + p_{10,5}X_5 + p_{10,4}X_4 + p_{10,3}X_3 + \\
&\qquad p_{10,2}X_2 + p_{10,1}X_1 + p_{10f}R_f \\
X_{11} &= p_{11,10}X_{10} + p_{11,9}X_9 + p_{11,8}X_8 + p_{11,7}X_7 + p_{11,6}X_6 + p_{11,5}X_5 + p_{11,4}X_4 + \\
&\qquad p_{11,3}X_3 + p_{11,2}X_2 + p_{11,1}X_1 + p_{11g}R_g
\end{aligned}
$$

For the model shown, which assumes that there are no unmeasured variables except the residual variables, that these residuals are uncorrelated, and that each dependent variable is directly linked to all other variables that precede it in the postulated causal sequence, the structural equations represent the basic normal equations of regression theory, and the path coefficients p_{ij} are the standardized partial regression coefficients [48]; for example, $p_{45} = \beta_{45.123}$ and $p_{64} = \beta_{64.1235}$.

The structural equations permit the correlation between any two variables in the model to be expressed in terms of the paths leading to the two variables from common antecedent variables as well as the direct path between them. This expression can be generalized in terms of what has been called the basic theorem of path analysis:

$$r_{ij} = \sum_k p_{ik} r_{kj}$$

where i and j denote two variables and the index k runs over all paths leading directly to X_i [49, 50]. For example,

$$r_{53} = p_{54}r_{43} + p_{53} + p_{52}r_{23} + p_{51}r_{13}$$

Table 1. Multiple Correlation Coefficients and Standardized Partial Regression Coefficients

Dependent variable	Independent variable	Multiple correlation coefficient	Standardized partial regression coefficient	Unstandardized partial regression coefficient
Mortality due to accidents, suicides, and cirrhosis of the liver	Hospital beds/ population ratio	.79	.08	1.722
	Per capita income		–.19	–.019
	Percent unemployment		.26	3.871
	Median education		–.04	–1.294
	Median age		.48	5.350
	Net migration		–.38	–1.404
	Percent agriculture		–.63	–2.663
	Percent urban		–.42	–.681
	Percent Spanish-American		–.03	–.064
	Percent nonwhite		.25	.953
Hospital beds/ population ratio	Per capita income	.76	.08	.040
	Percent unemployment		–.12	–8.172
	Median education		–.23	–38.674
	Median age		.56	29.785
	Net migration		–.36	–6.389
	Percent agriculture		–.24	–4.808
	Percent urban		.80	6.164
	Percent Spanish-American		.42	4.863
	Percent nonwhite		.49	8.980
Per capita income	Percent unemployment	.88	.06	9.444
	Median education		.80	279.495
	Median age		.28	31.060
	Net migration		–.04	–1.434
	Percent agriculture		–.06	–2.432
	Percent urban		.05	.734
	Percent Spanish-American		.05	1.299
	Percent nonwhite		–.11	–4.280
Percent unemployment	Median education	.88	–.37	–.881
	Median age		–.24	–.181
	Net migration		–.11	–.027
	Percent agriculture		–.50	–.014
	Percent urban		.10	.011
	Percent Spanish-American		.51	.083
	Percent nonwhite		.12	.032
Median education	Median age	.88	–.34	–.106
	Net migration		.06	.007
	Percent agriculture		–.15	–.018
	Percent urban		.21	.009
	Percent Spanish-American		–.80	–.054
	Percent nonwhite		–.58	–.063
Median age	Net migration	.74	.39	.129
	Percent agriculture		.26	.099
	Percent urban		–.16	–.024
	Percent Spanish-American		–.52	–.113
	Percent nonwhite		–.52	–.179
Net migration	Percent agriculture	.49	–.27	–.308
	Percent urban		.33	.142
	Percent Spanish-American		.09	.056
	Percent nonwhite		–.01	–.011

Each correlation can be expanded in terms of path coefficients. The result is

$$r_{53} = p_{53} + p_{54}p_{43} + p_{54}p_{42}p_{32} + p_{54}p_{42}p_{31}r_{12} + p_{54}p_{41}p_{32}r_{12} + p_{54}p_{41}p_{31} + p_{52}p_{32} + p_{52}p_{31}r_{21} + p_{51}p_{32}r_{12} + p_{51}p_{31}$$

The total effect of one variable on another is defined as the zero-order correlation between the two variables. In path analysis this total effect is partitioned into a direct effect, indirect effects through other intervening variables in the causal sequence, and a joint, or spurious, effect due to the mutual correlation of both variables with other variables in the model. For exogenous variables, this joint effect arises from the fact that part of the correlation between the dependent variable and the exogenous independent variable is due to the latter's correlation with the other exogenous variables in the model. When two endogenous variables are involved, the joint effect reflects the fact that the two variables have a number of common causes.

If p_{ij} is subtracted from both sides of the equation, the quantity $r_{ij} - p_{ij}$ is the expression for the total indirect effect of variable j on variable i. An examination of each of the terms on the right side of the equation provides valuable information regarding indirect effects, which in some instances may be far more important than the direct effect of one variable on another.

A formula for the residual path coefficient can also be derived from these equations [48, 49], as follows:

$$p_{ia} = \sqrt{1 - R^2}$$

where R^2 is the squared multiple correlation coefficient. This path coefficient measures the effect of all unmeasured variables not included in the model that cause variation in the endogenous variable.

Analysis and Results

Total Effects

As indicated above, once path coefficients have been estimated, analysis of the model begins with examination of the zero-order correlations among the variables, or the total effect of one variable on another. The intercorrelations among the variables in the model are shown in Table 2 (overleaf) for the 32 New Mexico counties; these correlation coefficients can be considered to be standardized zero-order regression coefficients.

For most of the endogenous variables included in the model, the total effect on the health status index is rather substantial. The zero-order correlations of the selected mortality rate with per capita income and with median number of school years completed by persons 25 years of age and over are both –.48. Median age and percent migration experienced between 1960 and 1969 have smaller negative correlations of –.17 and –.25, respectively, while percent unemployment has a high positive correlation with the index of health (.43).

Table 2. Intercorrelation Matrix (Zero-order Correlations)

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
X_1	1.00	–.52	.13	–.36	–.42	.30	–.28	–.01	–.16	–.40	–.20
X_2		1.00	–.49	–.04	.42	.14	.68	–.45	.59	.57	–.39
X_3			1.00	–.25	–.11	–.32	–.68	.75	–.52	–.24	.22
X_4				1.00	.05	–.46	–.18	.18	–.37	.15	.39
X_5					1.00	.24	.19	–.22	.20	.19	–.25
X_6						1.00	.18	–.55	.40	.21	–.17
X_7							1.00	–.71	.82	.20	–.48
X_8								1.00	–.65	–.19	.43
X_9									1.00	.24	–.48
X_{10}										1.00	.07
X_{11}											1.00
Mean	14.91	45.01	34.16	6.87	–15.23	24.08	9.95	5.65	2153.16	345.08	125.00
Standard deviation	12.29	32.01	21.31	13.40	13.95	4.63	1.45	3.49	508.13	246.52	51.62

X_1 = Percent of labor force in agriculture 1968
X_2 = Percent urban 1960
X_3 = Percent Spanish-American 1960
X_4 = Percent nonwhite 1960
X_5 = Net migration 1960–69
X_6 = Median age 1960
X_7 = Median education 1960
X_8 = Percent unemployed 1968
X_9 = Per capita income 1967
X_{10} = Hospital beds/100 000 population 1968
X_{11} = Mortality/100 000 population due to accidents, suicide, and cirrhosis of the liver 1968

The exogenous variables also evidence significant correlations with the health measure. Both the proportion of the labor force engaged in agriculture and the percentage of the population living in urban areas are negatively correlated with health status (–.20 and –.39, respectively). In addition, percent Spanish-American has a correlation of .22 and percent nonwhite a correlation of .39 with the health measure, indicating positive total effects of these two variables used to describe the ethnic composition of the population.

Direct and Indirect Effects

In Table 3 the total effects of each variable are partitioned into direct and indirect effects. Both standardized and unstandardized regression coefficients, shown in Table 1, were used in order to obtain a clear understanding of the effects of each variable on health. Examining first the variables most proximate to the measure of health, the hospital beds per 100 000 population ratio appears to have a small positive direct effect on the mortality index. For each additional bed per 100 000 population, the number of deaths per 100 000 increases by almost two. The positive effect of this variable is unexpected and difficult to account for, since it was hypothesized that health status would be higher and hence the mortality lower in counties that offer better health services. Most likely the model fails to include a variable that is highly correlated with both the bed ratio and mortality. The observed effect then would be spurious. This finding needs to be explored in future research.

Table 3. Contribution of Variables to Mortality Index of Health Status
(Figures in parentheses are based on *standardized* partial regression coefficients)

Variable	Total effect*	Direct effect†	Total indirect effect	Indirect effect through other variables	Joint or spurious effect
Hospital beds to population ratio	(.07) 1.46	(.08) 1.67	(−.01) −.21	. . .	(−.01) −.21
Per capita income	(−.48) −.05	(−.19) −.02	(−.29) −.03	(.01) 0	(−.30) −.03
Percent unemployment	(.43) 6.36	(.26) 3.84	(.17) 2.51	(−.02) −.30	(.19) 2.81
Median education	(−.48) −17.11	(−.04) −1.43	(−.44) −15.68	(−.25) −8.91	(−.19) −6.77
Median age	(−.17) −1.89	(.48) 5.34	(−.65) −7.23	(.04) .45	(−.69) −7.68
Net migration	(−.25) −.92	(−.38) −1.40	(.13) .48	(.13) .48	(0) 0
Percent agriculture	(−.20) −.84	(−.63) −2.64	(.43) 1.80	(.12) .50	(.31) 1.30
Percent urban	(−.39) −.63	(−.42) −.68	(.03) .05	(−.15) −.24	(.18) .29
Percent Spanish-American	(.22) .53	(−.03) −.08	(.25) .61	(.08) .20	(.17) .41
Percent nonwhite	(.39) 1.50	(.25) .96	(.14) .54	(−.01) −.04	(.15) .58

* Zero-order correlation.
† Direct-path coefficient.

Reexamination of the path coefficients in Table 1 linking social, economic, and demographic variables to this facility measure reveals a major long-term problem regarding location of beds, however. The hospital beds/population ratio is higher in those counties that contain large minority ethnic groups: path coefficients linking percent nonwhite and percent Spanish-American to this ratio are .49 and .42, respectively. This indicates that a 1 percent increase in the nonwhite population in a county increases the bed ratio by 9 beds/100 000; for a 1 percent increase in Spanish-Americans the increase is almost 5 beds/100 000. Moreover, the sign of the path coefficient linking migration to facilities is negative (−.36), indicating that the major hospital facilities in New Mexico are generally located in counties where out-migration exceeds in-migration—a finding of ominous import, since it suggests that a lag in the development of hospital facilities has occurred.

Per capita income appears to have a negligible effect on health. An increase of $100 in per capita income is required to reduce the mortality rate by 2 deaths per 100 000. While the effect of this variable on mortality rate is in the expected direction, its magnitude is somewhat smaller than expected.

Unemployment has a large direct effect on the mortality due to accidents, suicide, and cirrhosis of the liver. The path coefficient is .26, while the indirect effect through other variables is rather small. A 1 percent increase in unemployment results in 4 additional deaths per 100 000.

It is interesting to note that the zero-order correlation between unemployment and mortality (.43) overestimates the true effect of this variable. This results from the joint correlation of unemployment and mortality with percent Spanish-American, percent urban, and median education in particular. The resulting spurious effect of .19, then, accounts for almost all the indirect effect of this variable on health and a significant portion of its total effect.

The nature of the direct effect of this variable on the index mortality rates has important implications. These mortality rates appear to be significantly linked to unemployment even when the age structure, educational level, ethnic composition, and urban structure of the county are taken into account. This suggests that the debilitating effect of unemployment in areas where there are few job opportunities and little hope for improvement in economic status may result in an increased death rate due to these causes.

Education has a surprisingly small direct effect on these mortality rates (–.04), especially in the light of its high zero-order correlation of –.48. While part of this total effect is spurious (–.19) and is caused by the mutual correlation of education and mortality with common causes (in particular, percent Spanish-American and percent urban), at the same time indirect effects through other variables result in a large net indirect effect of –.25. The total of the direct effect and the indirect effect through other variables is such that the mortality rate declines by 10 deaths per 100 000 for each additional year of education completed. Education primarily affects economic variables in the model. As a county's educational level increases by one additional year, the unemployment rate decreases by almost 1 percent and per capita income increases by $279 (Table 1); unemployment in particular has a large direct effect on the mortality rate. This finding provides additional support for the thesis that the economic well-being of a county population is a major factor in accounting for differential mortality rates due to accidents, suicide, and cirrhosis of the liver.

As expected, age appears to play a significant role in mortality due to these causes. The direct effect of age on mortality is .48. As the age of the county's population rises by one year, the mortality rate increases by 5 deaths per 100 000. Indirect effects through other variables are negligible. In this case, the correlation of these two variables with antecedent variables in the causal chain act to suppress the true effect of age on mortality. Almost all the total indirect effect of –.65 is spurious. Since this effect is in the opposite direction from the direct effect of this variable, the total effect, or zero-order correlation, is only –.17. An examination of the zero-order correlation alone in this instance would result in a seriously erroneous conclusion regarding the true effect of this variable on health status.

Migration also partly accounts for differential mortality rates. The direct effect of this variable on the health status index is –.38, indicating that as out-migration exceeds in-migration the mortality rate increases. The direct effect in this instance accounts for virtually all the total effect of migration on health, since the net indirect effect of this variable through other variables is negligible. This finding is not too surprising, since social and economic conditions generally deteriorate as an area experiences out-migration over long periods of time, with a resultant impact on levels of health among the population remaining in the area. Additionally, migration is generally selective: the more able, healthier, better-educated individuals in a population are the most mobile and are likely to migrate to urban areas where job opportunities are greater and living conditions are generally better. The effect of this variable is thus as expected.

Counties with economies primarily based on agriculture also have much lower mortality rates. A 1 percent increase in the percent of the labor force engaged in agriculture results in a decrease in the mortality rate of about 3 deaths per 100 000. The more agrarian-based the county the lower the mortality rates due to accidents, suicide, and cirrhosis of the liver. An agriculturally based economy has other implications also, in that unemployment levels are generally lower, which results in a reduction in these rates of mortality. This indirect effect, however, is largely offset by the effect of agriculture on age: the median age of the population in agrarian-based counties is higher, which tends to increase mortality rates. The low zero-order correlation results from a large joint effect due to the correlation of percent agriculture with other exogenous variables, specifically, percent urban and percent nonwhite. These correlations result in a joint effect of .31, which masks the large direct effect of this variable on health status.

Urbanization of a county's population has a very slight effect on mortality rates. For every 1 percent increase in the proportion of the population living in urban areas there is a decrease of about 1 death in the mortality rate from accidents, suicides, and cirrhosis of the liver. Such a result might be expected because of the better living conditions and the general availability of health services in urban areas; however, the magnitude of this effect is rather small.

The ethnic composition of the county is another factor in determining the death rate, which increases as the nonwhite population increases. The path coefficient is .25, but some important indirect effects partly offset the direct effect of this variable. Nonwhites are younger than the population as a whole, consequently counties with a high proportion of nonwhites have a lower median age. This results in an indirect effect through age. Other indirect effects, however, largely cancel this effect. In particular, nonwhites are more poorly educated, which results in increased mortality. The net indirect effect through other variables, then, is only –.04.

While the direct effect of the percentage Spanish-American on the index mortality rate is only –.03, the indirect effect through other variables is .08. This variable appears to affect mortality primarily through its effect on age, education, and unemployment. Counties with a high proportion of Spanish-Americans have, on the whole, younger populations, and this factor has a large effect on the index mortality rate. This indirect effect is offset, however, by education and unemployment, since such counties have lower education levels and higher unemployment rates, resulting in an increase in mortality due to accidents, suicide, and cirrhosis of the liver. The net indirect effect of percent Spanish-American through other variables, then, is rather small, .08; and when this is combined with the direct effect of –.03, the net effect is only .05.

This marked difference in effect between the two ethnic variables may be the result of differences in the social control structure between the two ethnic groups. Other studies have shown that deviance, especially excessive use of alcohol, is lower among Spanish-Americans than among American Indians, who make up most of the nonwhite population of New Mexico [70, 71]. The lower rate among Spanish-Americans has been accounted for by their limited exposure to deviant role models and the existence of strong social sanctions resulting from a network that includes the Catholic Church, the family, and informal groups.

Summary and Conclusions

The model described here has been shown to be quite effective in predicting rates of mortality from accidents, suicide, and cirrhosis of the liver, with the model variables accounting for 64 percent of the variation in this health index.

More importantly, the model provides insights into the effects of population dynamics on a community's health status. The economic structure of a county was found to be a significant determinant of the mortality-rate health index. Mortality due to accidents, suicide, and cirrhosis of the liver is high in New Mexico counties that are not agriculturally based and have not developed other sectors of their economy such as manufacturing or retail and wholesale trade to provide employment for their populations. The resulting unemployment in these counties may account for many of the health problems experienced by their populations.

A second major set of factors has to do with social characteristics of the population, especially urbanization and migration. The death rate due to accidents, suicide, and cirrhosis of the liver declines as a direct effect of urbanization, with a further decrease in this mortality rate mediated through the younger age and higher education levels stemming from urbanization. Net migration, on the other hand, apparently results in a deterioration in the overall health status of the population when it is negative, although this finding rather unexpectedly does not appear to be mediated through education.

Finally, the demographic composition of the population, in particular age and ethnic composition, are important factors. The ethnic composition of the population affects health status, but in some unexpected ways. The one important direct effect is from the percentage of the population classified as nonwhite (mostly American Indians in New Mexico), with deaths due to accidents, suicide, and cirrhosis of the liver more frequent in counties that have large nonwhite populations. This effect is slightly ameliorated by the relative youthfulness of this ethnic group. The proportion of Spanish-Americans residing in the county, unlike the percent nonwhite, has little effect on this index of health status, perhaps because of the social structure peculiar to Spanish-American communities. Ethnic composition affects mortality also in indirect ways, through the lower educational levels among both Spanish-Americans and nonwhites and their higher levels of unemployment.

The path analysis of the data from New Mexico indicates that present hospital facilities are located in counties experiencing high mortality rates—counties that at the same time are losing population through out-migration, have large, poorly educated ethnic minority populations, and have limited employment opportunities with attendant high levels of unemployment. Counties with a more attractive economic base that are rapidly urbanizing under the influence of in-migration have much lower hospital beds/population ratios. These findings suggest that the existing hospital facilities in New Mexico may soon be inadequate to handle the burgeoning population of the more urban counties.

In conclusion, the model as well as the algorithms presented here are valuable in predicting the impact of structural changes in social systems brought about by natural causes as well as by planned intervention. The approach elaborated here provides a means of constructing a causal model from a limited number of variables postulated to be related to health measures and to be interrelated among themselves. Parameters of the model can be estimated from data and the model evaluated. According to the results, the model can be either reformulated or used to make predicitions. These predictions are important to an understanding of the complex social processes that occur in the social system as structural changes occur, as well as in predicting the direct and indirect outcomes of such programs.

REFERENCES

1. Anderson, J. G. Social Indicators and Second-order Consequences: Measuring the Impact of Innovative Health and Medical Care Delivery Systems. Paper presented at the 2d National Conference on Systems Research for Improving the Delivery of Health Care Services, Hospital and Health Services Division, AIIE, Denver, Feb. 4, 1971.
2. Fanshel, S. and J. W. Bush. A health status index and its application to health services outcomes. *Op. Res.* 18:1021, 1970.
3. *Conceptual Problems in Developing an Index of Health.* Vital and Health Statistics, Series 2, No. 17. Washington: National Center for Health Statistics, 1966.
4. *An Index of Health: Mathematical Models.* Vital and Health Statistics, Series 2, No. 5. Washington: National Center for Health Statistics, 1965.
5. Sanders, B. S. Measuring community health levels. *Am. J. Pub. Health* 54:1063, 1964.
6. Packer, A. H. Applying cost-effectiveness concepts to the community health system. *Op. Res.* 16:227, 1968.
7. Flagle, C. D. Operations research in the health services. *Op. Res.* 10:591, 1962.
8. Forrester, J. W. Counterintuitive behavior of social systems. *Technol. Rev.* 73:53, 1971.
9. Forrester, J. W. Systems analysis as a tool for urban planning. *IEEE Spectrum* 8:48, 1971.
10. Madison, D. L. Organized health care and the poor. *Med. Care Rev.* 26:1, 1969.

11. Turner, W. W. Satellite hospitals bring health care to suburbanites. *Mod. Hosp.* 99:92, 1962.
12. Cardwell, R. L., M. G. Reed, and M. Shain. *Hospital Utilization in a Metropolitan Area.* Chicago: Hospital Planning Council for Metropolitan Chicago, 1964.
13. Shannon, G. W., R. L. Bashshur, and C. A. Metzner. The concept of distance as a factor in accessibility and utilization of health care. *Med. Care Rev.* 26:143, 1969.
14. Weiss, J. E. and M. R. Greenlick. Determinants of medical care utilization: The effect of social class and distance on contacts with the medical care system. *Med. Care* 8:456, 1970.
15. Weiss, J. E., M. R. Greenlick, and J. F. Jones. Determinants of medical care utilization: The impact of ecological factors. Paper presented at the 98th Annual Meeting of the American Public Health Association, Houston, Oct. 26–30, 1970.
16. Bailey, N. T. J. Operations research in medicine. *Op. Res. Quart.* 3:24, 1952.
17. Bailey, N. T. J. Operational research in hospital planning and design. *Op. Res. Quart.* 8:149, 1957.
18. Bailey, N. T. J. *The Mathematical Approach to Biology and Medicine.* New York: Wiley, 1967.
19. Flagle, C. D. Operations research in a hospital. In C. D. Flagle, W. Huggins, and R. Roy (eds.), *Operations Research and Systems Engineering*, p. 763. Baltimore: Johns Hopkins Press, 1960.
20. Flagle, C. D. and J. P. Young. Application of operations research and industrial engineering to problems of health services, hospitals, and public health. *J. Ind. Eng.* 17:609, 1966.
21. Middlehoven, W. Analysis and Reorganization of a Central Supply Delivery System. Master's thesis, Department of Operations Research and Industrial Engineering, Johns Hopkins University, 1964.
22. Balintfy, J. L. Menu planning by computer. *Communications ACM* 7:255, 1964.
23. Gue, R. L. and J. C. Liggett. Mathematical programming models for hospital menu planning. *J. Ind. Eng.* 17:395 Aug. 1966.
24. Hsieh, R. K. C. A Study of Linen Processing and Distribution in a Hospital. Master's thesis, Department of Industrial Engineering, Johns Hopkins University, 1961.
25. Barker, K. N. and W. M. Heller. The development of a centralized unit dose dispensing system. *Am. J. Hosp. Pharm.* 20:568, November 1963.
26. Hsieh, R. K. C. Evaluation of Hospital Medication Systems. Doctoral dissertation, Department of Public Health Administration, Johns Hopkins University, 1966.
27. Beenhakker, H. L. Multiple correlation: A technique for prediction of future hospital bed needs. *Op. Res.* 11:824, 1963.
28. Feldstein, P. J. and J. J. German. Predicting hospital utilization: An evaluation of three approaches. *Inquiry* 2:13, 1965.
29. Rosenthal, G. D. Factors affecting the utilization of short term general hospitals. *Am. J. Pub. Health* 55:1734, 1965.
30. Balintfy, J. L. A stochastic model for the analysis and prediction of admissions and discharges in hospitals. In C. W. Churchman and M. Verhulst (eds.), *Management Sciences: Models and Techniques*, Vol. 2, p. 288. New York: Pergamon, 1960.
31. Young, J. P. Stabilization of inpatient bed occupancy through control of admissions. *Hospitals, J.A.H.A.* 39:41, 1965.
32. Young, J. P. Administrative control of multiple channel queueing systems with parallel input streams. *Op. Res.* 14:145, 1966.
33. Flagle, C. D. et al. *Analysis of Congestion in an Outpatient Department.* Final report, USPHS Grant W-96. Department of Operations Research, Johns Hopkins Hospital, 1962.
34. Welch, J. D. Appointment systems in hospital outpatient departments. *Op. Res. Quart.* 15:224, 1965.
35. Fry, J. Appointments in general practice. *Op. Res. Quart.* 15:233, 1965.
36. Jackson, R. R. P. Design of an appointment system. *Op. Res. Quart.* 15:219, 1965.
37. Bailey, N. T. J. A study of queues in hospital outpatient departments with special reference to waiting times. *J. Roy. Statist. Soc.* (Ser. B) 14:185, 1952.
38. Welch, J. D. and N. T. J. Bailey. Appointment systems in hospital outpatient departments. *Lancet* 1:1105, 1952.

39. Fetter, R. B. and J. D. Thompson. The simulation of hospital systems. *Op. Res.* 13:689, 1965.
40. Katz, J. H. Simulation of out-patient appointment systems. *Communications ACM* 12:215, 1969.
41. Rising, E. J. et al. Simulation of a University Health Service Outpatient Department. Paper presented at the American College Health Association Meeting, San Francisco, April 27–30, 1971.
42. Rosenhead, J. V. Experimental simulation of a social system. *Op. Res. Quart.* 19:289, 1968.
43. Colley, J. L. et al. *Computer Simulation Models of Nursing Homes.* Research Triangle Institute, Chapel Hill, N.C., 1967.
44. Gentry, J. D., F. D. Kennedy, and A. H. Packer. *The Development of a Simulation Model of a Community Health Service System.* Vol. II, *The Health Service System.* Research Triangle Institute, Chapel Hill, N.C., 1968.
45. Kennedy, F. D. *The Development of a Simulation Model of a Community Health Service System.* Vol. I, *Introduction and Summary.* Research Triangle Institute, Chapel Hill, N.C., 1968.
46. Kennedy, F. D. and M. B. Woodside. *The Development of a Simulation Model of a Community Health Service System.* Vol. III, *The Maternal and Infant Care Simulation Model.* Research Triangle Institute, Chapel Hill, N.C., 1968.
47. Milly, H. and L. J. Pocinki. *A Computer Simulation Model for Evaluation of the Health Care Delivery System.* GEOMET, Inc., Rockville, Md., 1970.
48. Duncan, O. D. Path analysis: Sociological examples. *Am. J. Sociol.* 72:1, 1966.
49. Land, K. C. Principles of path analysis. In E. F. Borgatta (ed.), *Sociological Methodology 1969*, p. 3. San Francisco: Jossey Bass, 1969.
50. Blalock, H. M. *Causal Inferences in Nonexperimental Research.* Chapel Hill: University of North Carolina Press, 1964.
51. Blalock, H. M. *Causal Models in the Social Sciences.* Chicago: Aldine, 1971.
52. Wright, S. The method of path coefficients. *Ann. Math. Statist.* 5:161, 1934.
53. Wright, S. The interpretation of multivariate systems. In O. Kempthorne et al. (eds.), *Statistics and Mathematics in Biology*, p. 11. Ames, Iowa: Iowa State University Press, 1954.
54. Wright, S. Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics* 16:189, 1960.
55. Li, C. C. The concept of path coefficient and its impact on population genetics. *Biometrics* 12:190, 1956.
56. Wold, H. O. Causality and econometrics. *Econometrica* 22:162, 1954.
57. Wold, H. O. and L. Jureen. *Demand Analysis.* New York: Wiley, 1953.
58. *Acute Conditions: Incidence and Associated Disability, United States, July 1966–June 1967.* Vital and Health Statistics, Series 10, No. 44. Washington: National Center for Health Statistics, March 1968.
59. *Limitation of Activity and Mobility Due to Chronic Conditions, United States, July 1965–June 1966.* Vital and Health Statistics, Series 10, No. 45. Washington: National Center for Health Statistics, May 1968.
60. *Disability Days, United States, July 1965–June 1966.* Vital and Health Statistics, Series 10, No. 47. Washington: National Center for Health Statistics, October 1968.
61. White, E. L. A graphic presentation on age and income differentials. *Inquiry* 5:18, 1968.
62. Richardson, W. C. Poverty, illness, and use of health services in the United States. *Hospitals, J.A.H.A.* 43:34, 1969.
63. *Volume of Physician Visits, United States, July 1966–June 1967.* Vital and Health Statistics, Series 10, No. 49. Washington: National Center for Health Statistics, November 1968.
64. *Persons Hospitalized, by Number of Hospital Episodes and Days in a Year, United States, July 1965–June 1966.* Vital and Health Statistics, Series 10, No. 50. Washington: National Center for Health Statistics, February 1969.
65. Adams, B. N. *Kinship in an Urban Setting.* Chicago: Markham, 1968.
66. Gore, M. S. *Urbanism and Family Change.* New York: Humanities Press, 1968.
67. Gibbs, J. P. and W. T. Martin. Urbanization, technology, and the division of labor:

International patterns. *Am. Sociol. Rev.* 27:667, 1962.
68. *Health Characteristics, by Geographic Region, Large Metropolitan Area, and Other Places of Residence, United States, July 1963–June 1965.* Vital and Health Statistics, Series 10, No. 36. Washington: National Center for Health Statistics, April 1967.
69. Colorado Commission on Spanish-surnamed Citizens. *The Status of Spanish-surnamed Citizens in Colorado.* Department of Political Science, Colorado State College, Greeley, Colo., 1967.
70. Jessor, R. Toward a social psychology of excessive alcohol-use. In C. R. Snyder and D. R. Schweitzer (eds.), *Proceedings, Research Sociologists' Conference on Alcohol Problems,* p. 59. Southern Illinois University, Carbondale, 1964.
71. Jessor, R. et al. *Society, Personality and Deviant Behavior.* New York: Holt, Rinehart and Winston, 1968.